REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
28/02/2010

2. REPORT TYPE
Final Performance Report

3. DATES COVERED *(From - To)*
01/05/2007 - 30/04/2010

4. TITLE AND SUBTITLE
Adaptive problem solving and decision making in complex and changing environments: The role of understanding and explanation

5a. CONTRACT NUMBER
FA9550-07-1-0147

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
John E. Hummel (PI)
Brian H. Ross (Co-PI)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Psychology
University of Illinois
203 E. Daniel St.
Champaign, IL 61820

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
875 N. Randolph Street
Suite 324, Room 3112
Arlington VA 22203-1768

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0371

12. DISTRIBUTION/AVAILABILITY STATEMENT
A- Approve For Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The ability to generate explanations plays a central role in human cognition and is essential for intelligent problem solving and decision making. Generating explanations requires a deep understanding of the domain and tremendous flexibility in the way concepts are accessed, combined and used. Together, the joint requirements of deep understanding and flexibility in conceptual access and use constitute challenging design requirements for a model of explanation. The PIs developed a systematic program of computational modeling to elucidate the mental representations and processes underlying the generation of explanations in the service of problem solving and decision making. They developed and implemented a detailed process model of explanation—one that is capable of flexibly using its knowledge in the service of explaining novel explanada—and demonstrated that it provides a good qualitative account of patterns of explanations generated by human adults in laboratory settings. The same principles the model uses to generate explanations (i.e., reasoning "backward" to infer causes of known effects) are also applicable to planning and problem solving (i.e., reasoning "forward" to infer useful actions for problem solving).

15. SUBJECT TERMS
Explanation, human reasoning, computational modeling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			John E. Hummel
					19b. TELEPHONE NUMBER *(Include area code)* (217) 265-6090

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

To: technicalreports@afosr.af.mil
Subject: Final Progress Statement to Dr. Jun Zhang

Contract/Grant Title: Adaptive problem solving and decision making in complex and changing environments: The role of understanding and explanation
Contract/Grant #: FA9550-07-1-0147

Reporting Period: 1 May, 2007 to 30 April, 2010 (with a no-cost extension to 30 October, 2010)

Executive Summary

The ability to generate explanations plays a central role in human cognition and is essential for intelligent problem solving and decision making. Generating explanations requires a deep understanding of the domain and tremendous flexibility in the way concepts are accessed, combined and used. Together, the joint requirements of deep understanding and flexibility in conceptual access and use constitute challenging design requirements for a model of explanation.

The PIs developed a systematic program of computational modeling to elucidate the mental representations and processes underlying the generation of explanations in the service of problem solving and decision making. The guiding insight underlying this effort was that the process of explanation-generation shares many of the same computational requirements as the process of analogy-making. In particular, both depend on the flexible use of rich systems of relational knowledge. Accordingly, the PIs' starting point for this modeling effort was Hummel and Holyoak's (1997, 2003) LISA model of analogy, analogical inference and schema induction.

Explanation differs from analogy in several important respects, however, requiring the PIs to expand the LISA model in a number of important directions. First, whereas analogy-making is an extremely content-general cognitive process (in the sense that people can make analogies about virtually anything, with no particular set of relations or content knowledge being privileged over others), *causal* relations play a privileged role in explanation. Explanations can incorporate all kinds of relational knowledge (e.g., invoking relations between fuel lines and fuel injectors in an explanation for why a car won't start) but all these relations are tied together into explanations by higher-order causal relations (e.g., the fuel line causes the fuel to flow to the fuel injector). Moreover, these causal relations seem cognitively privileged over other relations in the sense that they guide the explanation process, structure the resulting explanations and seem to do so in a way that is more implicit (i.e., less taxing of working memory resources) than the explicit, lower-level relations they structure.

Accordingly, the PIs developed a novel approach to representing causal relations in the LISA model. LISA is an artificial neural network that represents propositional knowledge (e.g., *mixes* (fuel-injector, fuel, oxygen)) in a hierarchy of neuron-like units. At the bottom of the hierarchy, *semantic* units represent the semantic features of objects (such as fuel-injector, fuel and oxygen) and relational roles (such as the roles of the *mixes* relation). Together, the semantic units represent objects and relational roles in a distributed fashion, explicitly capturing what different objects and roles have in common and how they differ. Above the distributed semantic units, localist *object* and *role* units represent objects and relational roles in a local fashion, sharing bi-directional excitatory connections with the corresponding semantic units. Localist *role-binding* units (also called *sub-proposition*, or *SP*, units) represent bindings of objects (or whole propositions) to relational roles. For example, to represent mixes (fuel-injector, fuel,

20120918107

oxygen), one SP would represent the binding of the first role of *mixes* to fuel-injector, a second would bind the second role to fuel and a third the third role to oxygen. These units share bi-directional excitatory connections with the object (or proposition) and role units to which they refer. Finally, at the top of the representational hierarchy in LISA, *proposition*, or P, units bind sub-propositions into full propositions by sharing excitatory connections with the corresponding SP units. When a proposition enters working memory (i.e., when it becomes active) the SPs representing its various role bindings fire out of synchrony with one another, causing the representations of its constituent roles to fire *in* synchrony with their arguments and *out* of synchrony with one another (e.g., the first role of *mixes* would fire in synchrony with fuel-injector and out of synchrony with the second and third roles and their arguments).

This approach to knowledge representation is capable of representing causal relations as explicit relations (i.e., with the first role of *causes* firing in synchrony with the cause, the second role firing in synchrony with the effect and the two roles firing out of synchrony with one another). However, Hummel and Holyoak (1997, 2003) showed that LISA's approach to knowledge representation provides an excellent account of the limitations of working memory (WM). As such, representing causal relations in this way would consume WM resources. (Although they agree that causal relations can be represented and reasoned about in this way, the PIs also believe that for the purposes of explanation-generation, causal relations can also be represented in a less WM-consuming fashion.) Accordingly, the PIs proposed a fourth kind of unit for LISA, the *group* unit. Group units "sit above" P (proposition) units n the representational hierarchy and share excitatory connections with the proposition(s) they group. The PIs used group units to represent propositions that collectively form a cause and those that form an effect. Finally, cause and effect groups are connected by higher-level group units that link causes to their effects. This approach allows LISA to represent causal relations explicitly but without incurring any additional load on WM. They also allow causal relations to form a natural basis for organizing knowledge in long-term memory (LTM) and for controlling both retrieval from LTM and the flow of control in reasoning and explanation-generation more generally (Hummel, Devnich & Landy, 2008, Hummel & Landy, 2009).

In addition to the privileged role of causal relations, another important way in which explanation differs from analogy is that it often draws on multiple sources of knowledge from LTM. Analogy, by contrast, is typically conceived of as drawing on only a single source from LTM. This difference is extremely important as it speaks to one of the fundamental constraints that is broadly agreed to make the process of analogy-making possible in the first place: The one-to-one mapping constraint. As universally conceived in the analogy literature, analogy is the process of using a familiar *source* situation to reason about a novel *target* situation. The classic example is the analogy between the structure of the solar system and the Rutherford mode of the atom. In this analogy, the nucleus of the atom corresponds to the sun and the electrons to the planets. Like the planets, the electrons "orbit" the nucleus. (Although this model is now known to be incorrect, it was nonetheless at one time useful.) Armed with this analogy, and with the knowledge that a force (namely, gravity) causes the planets to orbit the sun, one can make the analogical inference that some force must cause the electrons to orbit the nucleus. This analogical inference depends critically on the *mapping* of the nucleus to the sun, the electrons to the planets and "orbiting" (in the case of the electrons) to orbiting in the case of the planets: The process of finding this mapping (i.e., the set of *relational* correspondences between elements of the two situations) is the very heart and soul of analogy-making. And it is universally agreed in the analogy literature that analogical mapping honors a one-to-one mapping constraint: Each object and relational role in one situation may correspond to at most one object or relational role in the other (e.g., if the electrons correspond to the planets then they cannot also correspond to the sun). Without this constraint, analogical mapping would be hopelessly ill-posed.

Explanation-generation, by contrast, typically calls on multiple sources of knowledge. Consider a simple example from Patalano, Chin-Parker and Ross (2006). These researchers gave subjects a problem of the form, "In the population as a whole, people tend to prefer Pepsi to Coke about as often as they prefer Coke to Pepsi. However, it turns out that ministers tend to prefer Pepsi over Coke," and asked them to generate an explanation for this "fact". Their subjects had no difficulty doing so, and all their explanations drew on multiple sources of knowledge, including knowledge about ministers, Coke and Pepsi both as products and corporations, and about things such as peoples' generic product preferences. For example, one common explanation took the form, "Well, Coke used to contain cocaine, and cocaine is illegal, so maybe ministers object to the Coke corporation on moral grounds." Note that this explanation integrates knowledge of ministers, the history of the Coke corporation, the legal status of cocaine and the kinds of things that might lead an individual to prefer one company's product over another's. Integrating multiple sources of knowledge presents many complexities for a model of explanation-generation, not least of which is that it requires the reasoner/model to violate the one-to-one mapping constraint: In the context of a "minister" schema (i.e., one's generic knowledge about what ministers are typically like) the minister in the stated problem corresponds (i.e., maps) to the minister in the schema; but in a "product preference" schema, that same minister corresponds to "generic-product-preferring-person".

Generalizing LISA's algorithm for analogy-making to the problem of explanation-generation thus required the PIs to find a way to violate the one-to-one mapping constraint without rendering the analogical mapping problem fundamentally ill-posed. They did so by implementing a procedure, controlled by the very group units that represent causal relations, that iteratively retrieves relevant knowledge from LTM, maps that knowledge onto the explanandum (i.e., that which is to be explained), uses it to make inferences about the explanandum, then "forgets" the mappings that drove those inferences and repeats the retrieve-map-infer cycle. The resulting algorithm provides a good qualitative account of the kinds of explanations subjects generate in the laboratory (including those observed by Patalano et al., 2006; see Hummel et al., 2008; Hummel & Landy, 2009).

In addition to the LISA-based process model of explanation, the PIs also developed a computational-theory level model of explanation, ERIC (Explanatory Reasoning for Inductive Confidence), that combines Bayesian reasoning with analogy to generate explanations in the service of updating its inductive confidence in its beliefs. The resulting model, which the PIs are still developing, does an excellent job accounting for a large body of effects in the literature on inductive confidence (Landy & Hummel, 2009, 2010). The PIs have also completed numerous empirical projects related to explanation and related problems, such as relational reasoning and concept acquisition. At least 17 papers, chapters and conference proceedings, as well as numerous presentations at scientific meetings, have been credited to the grant. The researchers supported by this grant included John Hummel (PI), Brian Ross (Co-PI), David Landy (post-doc), Erin Jones, Wookyoung Jung, Eric Taylor (graduate students), Pamela Glosson and Robert Weisshappel (undergraduate research assistants).

References

Hummel, J. E., Landy, D. H., & Devnich, D. (2008). Toward a process model of explanation with implications for the binding problem. In *Naturally Inspired AI: Papers from the AAAI Fall Symposium. Technical Report FS-08-06*, 79-86.

Landy, D. H. & Hummel, J. E. (2010). Explanatory reasoning for inductive confidence. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*.

Patalano, A. L., Chin-Parker, S., & Ross, B. H. (2006). The importance of being coherent: Category coherence, cross-classification, and reasoning. *Journal of Memory and Language, 54*, 407-424.

Archival publications (published) during reporting period:

Taylor, E. T., Landy, D. & Ross, B. H. (revision under review). The benefits of explanations in simple binary prediction tasks.

Hummel, J. E. (in press). Object recognition. In D. Reisberg (Ed.) *Oxford Handbook of Cognitive Psychology*. Oxford, England: Oxford University Press.

Hummel, J. E. (in press). The proper treatment of symbols in a neural architecture. In S. Levy and R. Gayler (Eds.) *Compositional Connectionism in Cognitive Science II: The Localist/Distributed Dimension*.

Doumas, L. L. A., & Hummel, J. E. (2010). A computational account of the development of the generalization of shape information. *Cognitive Science*, *34*, 698 – 712.

Hummel, J. E. (2010). Symbolic vs. associative learning. *Cognitive Science*, *34*, 958-965.

Landy, D. H. & Hummel, J. E. (2010). Explanatory reasoning for inductive confidence. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*.

Hummel, J. E., & Landy, D. H. (2009). From analogy to explanation: Relaxing the 1:1 mapping constraint… Very carefully. In B. Kokinov, K. Holyoak and D. Gentner (Eds.) *New Frontiers in Analogy Research: Proceedings of the Second International Conference on Analogy*. Sofia, Bulgaria.

Jung, W., & Hummel, J. E. (2009). Learning probabilistic relational categories. In B. Kokinov, K. Holyoak and D. Gentner (Eds.) *New Frontiers in Analogy Research: Proceedings of the Second International Conference on Analogy*. Sofia, Bulgaria.

Jung, W. & Hummel, J. E. (2009). Probabilistic relational categories are learnable as long as you don't know you're learning probabilistic relational categories. *Proceedings of The 31st Annual Conference of the Cognitive Science Society*.

Knowlton, B. J., McAuliffe, S. P., Coelho, C. J., & Hummel, J. E. (2009). Visual priming of inverted and rotated objects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*, 837-848.

Landy, D., & Goldstone, R. L. (2009). Pushing Symbols. *Proceedings of The 31st Annual Conference of the Cognitive Science Society*.

Penn, D. C., Cheng, P. W., Holyoak, K, J., Hummel, J. E., & Povinelli, D. J. (2009). There's more to thinking than propositions. Commentary on Mitchell et al., "The propositional nature of human associative learning," *Behavioral and Brain Sciences*, *32*, 221-223.

Taylor, E. G., & Hummel, J. E. (2009). Finding similarity in a model of relational reasoning. *Cognitive Systems Research*, *10*, 229-239.

Taylor, E.G., Landy, D.H., & Ross, B. H. (2009). The role of explanation in very simple tasks. *Proceedings of The 31st Annual Conference of the Cognitive Science Society*.

Holyoak, K. J., & Hummel, J. E. (2008). No way to start a space program. Commentary on R. Leech, D. Mareschal and R. P. Cooper, "Analogy as relational priming: A developmental and computational perspective on the origins of a complex skill," *Behavioral and Brain Sciences*, *31*, 388-389.

Hummel, J. E., Landy, D. H., & Devnich, D. (2008). Toward a process model of explanation with implications for the binding problem. In *Naturally Inspired AI: Papers from the AAAI Fall Symposium. Technical Report FS-08-06*, 79-86.

Landy, D., Jones, E., & Hummel, J. E. (2008). Why spatial-numeric associations aren't evidence for a mental number line. Proceedings of *The 30th Annual Conference of the Cognitive Science Society*. Washington, D.C.